NPS55-86-001

NAVAL POSTGRADUATE SCHOOL

Monterey, California

USE OF COLOR IN DIFFERENTIATING
FACTOR LEVELS IN SIMULATION OUTPUT
(PRELIMINARY REPORT)

P. A. W. LEWIS

JANUARY 1986

Prepared for:

Office of Naval Research
Arlington, VA 22217

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Admiral R. H. Shumaker
Superintendent

D. A. Schrady
Provost

This report was prepared for Office of Naval Research: P. A. W. Lewis

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: Unclassified

1b RESTRICTIVE MARKINGS:

2a SECURITY CLASSIFICATION AUTHORITY:

2b DECLASSIFICATION / DOWNGRADING SCHEDULE:

3 DISTRIBUTION / AVAILABILITY OF REPORT: Approved for public release; distribution is unlimited

4 PERFORMING ORGANIZATION REPORT NUMBER(S): NPS55-86-001

5 MONITORING ORGANIZATION REPORT NUMBER(S):

6a NAME OF PERFORMING ORGANIZATION: Naval Postgraduate School

6b OFFICE SYMBOL (If applicable):

7a NAME OF MONITORING ORGANIZATION: Office of Naval Research

6c ADDRESS (City, State, and ZIP Code): Monterey, California 93943-5100

7b ADDRESS (City, State, and ZIP Code): Arlington, VA 22217

8a NAME OF FUNDING / SPONSORING ORGANIZATION:

8b OFFICE SYMBOL (If applicable):

9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N0001486AF00001

8c ADDRESS (City, State, and ZIP Code):

10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61153N	RR014-05-01	4114469-01	WR24018

11 TITLE (Include Security Classification): USE OF COLOR IN DIFFERENTIATING FACTOR LEVELS IN SIMULATION OUTPUT (PRELIMINARY REPORT)

12 PERSONAL AUTHOR(S): Lewis, P. A. W.

13a TYPE OF REPORT: Technical

13b TIME COVERED: FROM ____ TO ____

14 DATE OF REPORT (Year, Month, Day): 1986 January

15 PAGE COUNT: 10

16 SUPPLEMENTARY NOTATION:

17 COSATI CODES

FIELD	GROUP	SUB-GROUP

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Simulation, Color Graphics, Sampling Experiment, Multi-factor Simulation, Multiserver Queue

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

A simulation is essentially a multi-factor statistical sampling experiment through which approximate answers to questions about some aspect of a system or statistic are obtained. Unfortunately the multi-factor aspect of the simulation is usually downplayed because of the difficulties of organizing and displaying the output as a function of various factors. Some graphical procedures using color are suggested for assessing the effect of the factors on the output. This is done in the context of an example of a multiserver queue. Classical analysis of variance techniques are usually not appropriate for this analysis because the data is non-normal and the mean is seldom an adequate or complete characterization of the output.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION:

22a NAME OF RESPONSIBLE INDIVIDUAL: P. A. W. Lewis

22b TELEPHONE (Include Area Code): (408)646-2283

22c OFFICE SYMBOL: Code 55Lw

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

USE OF COLOR IN DIFFERENTIATING FACTOR LEVELS IN SIMULATION OUTPUT (PRELIMINARY REPORT)

by

Professor Peter A. W. Lewis
Department of Operations Research
Naval Postgraduate School
Monterey, CA 93943-5100

ABSTRACT

A simulation is essentially a multi-factor statistical sampling experiment through which approximate answers to questions about some aspect of a system or statistic are obtained. Unfortunately, the multi-factor aspect of the simulation is usually downplayed because of the difficulties of organizing and displaying the output as a function of various factors. Some graphical procedures using color are suggested for assessing the effect of the factors on the output. This is done in the context of an example of a multiserver queue. Classical analysis of variance techniques are usually not appropriate for this analysis because the data is non-normal and the mean is seldom an adequate or complete characterization of the output.

INTRODUCTION

Simulation is a widely used tool for finding answers to problems which involve random elements in fields as diverse as Statistics, Physics, Operations Analysis, and Engineering. It is essentially a multi-factor statistical sampling experiment which, with a model, is performed on a digital computer. Moreover, the answers obtained from the simulation are functions of the various factors (attributes, independent variables, concomitants) in the problem.

A very simple example of such a problem is the multiserver queue (G/G/m) which occurs in banks, supermarkets, barbershops, etc. In this queueing system, arriving customers are served in order of arrival, with one of two queueing disciplines. The first discipline is where customers form a **single line** and are called for service from the head of that line as servers become free. The second queueing discipline is where customers join the separate queues which form in front of each server and stay there until served. As an idealization of this second discipline, an arriving customer joins the shortest and presumably the **fastest service** line or picks, with equal probability, from lines which have the tied, fewest number of customers awaiting service. No jumping from server to server is allowed.

The rough justification for the **single line** discipline, denoted as SL, over the shortest line, **fastest service** queueing discipline (FS) is that an arriving customer will not get stuck behind a customer who is already in service and who has an extremely long service time. However both disciplines are commonly used, and the question to be answered is which discipline is 'best'. Of course this question is unanswerable as posed until the proper quantification of the problem is decided upon. Thus, is it adequate to compare the mean waiting times in a stationary queue, or should

one compare the probabilities of not having to wait on arrival at the queue or the values of, say, the .99 quantile of the waiting time distributions? (The α-quantile of the distribution of a random variable, W, is the value, w_α, such that the probability of W being less than w_α is α).

Furthermore, even if one of the above quantifications is decided upon and its value is compared for the two service disciplines, the answer to the question as to which service discipline is best may depend on the level of several other factors, as follows:

a. The traffic intensity is, from what theory exists, a definite factor here. In this queueing problem the traffic intensity is given by the ratio of the expected service time, E(S), divided by the product of the number of servers, m, and the expected interarrival time of customers, E(A). This traffic intensity factor, t, must be less than 1 for a stable queue to exist and has a continuum of possible values between 0 and 1. Three values of t will be used in the subsequent analysis, 0.30, 0.60 and 0.75.

b. The number of servers, m, is a factor which may influence the output and must clearly be greater than one for the distinction between the **single line** (SL) and the **shortest line, fastest service**, (FS) disciplines to have any meaning. Three values of m will be used in the subsequent analysis, 3, 5, and 10

c. The distribution of service times (in G/G/m queues assumed to be independent and identically distributed) and the distribution of interarrival times (also assumed to be independent and identically distributed) are also factors in this queueing situation. Because of the complexity of these factors only four combinations of distributions will be used in the subsequent analysis:

   A - very **variable** service; very **variable** interarrival;
   B - very **regular** service; very **variable** interarrival;
   C - very **variable** service; very **regular** interarrival;
   D - very **regular** service; very **regular** interarrival:

   This rough categorization will be more specifically detailed later for the purpose of implementing the simulation; the categorization represents an attempt to cut down a complex factor to manageable proportions based on experience. A rough guess would be that if the single line queueing discipline is effective in cutting down, in some sense, the waiting times of customers, it will be in case A, where very variable or skewed service and interarrival times are encountered. In that case not only are long service times relatively common, but also customers may occasionally arrive in quick succession.

With three levels of the traffic intensity factor t, three values of m, the number of servers, four combinations of interarrival and service time distributions and two queueing disciplines, giving 72 factor combinations, a

simulation experiment with a fixed number of replications at each factor level is technically handleable by a four-way analysis of variance. The reason that this is not appropriate and that a graphical summary of the output is necessary is that by comparison of the means of the data are not adequate and the data is not normally distributed. In fact the waiting times will generally be very positively skewed with a discrete component representing the probability that a customers waiting time is exactly zero. Further complicating the standard anslysis is the fact that the relationship between the factors will not be linear. For example, in many queueing situations the mean waiting time, E(W), is proportional to the reciprocal of $(1-t)^2$.

An alternative to a blind and perhaps inappropriate application of analysis of variance techniques is to first look at the simulation output data in order to perform an initial, exploratory analysis. This may show immediately the salient points of the simulation, or suggest further formal analysis after, for example, a transformation of the data.

The problem of differentiating four factors graphically is not simple, and in this paper we will attempt to obtain as concise and as compact a summary of the output of the simulation experiment as possible. This will be done graphically by using multiple X-Y plots, with additional coding within each plot obtained by color, spacing, and line-type.

THE DATA FROM THE G/G/m QUEUE

To complete the specification of the multiple server queue discussed above it was assumed that the individual service times were Gamma distributed, as were the independent interarrival times.

To obtain 'regular' service, it was assumed that the shape parameter, k, in the Gamma distribution took on the value 5.0; to obtain very 'variable' i.e. 'highly postively skewed' service times it was assumed that the shape parameter k took on the value 0.5. Similar conventions were made for the interarrival times.

The Gamma assumption was made for computational convenience; another possibility is to assume that the 'variable' distribution is obtained by a mixture of two exponential random variables. In that case analytical solutions can, in principle, be obtained for parts of the problem by the method of (parallel) stages (Kleinrock 1975).

For each set of the 72 factor levels, the queue was simulated out to the 5,000th arrival. This was repeated independently 440 times to obtain a vector of i.i.d realizations of waiting times of the 5000th customer. Since the traffic intensity was chosen to have the levels 0.30, 0.60 and 0.75, it can be assumed that the samples represent the stationary waiting time in the queue. It could be argued that a higher level of t, say 0.90, should also be used. However, not only would it then be necessary to go out even beyond the 5000th customer in the sample path of the queue to assure stationary, but also the already complex data handling problem with the 72 waiting time vectors would be further complicated. Also, as will be seen, the essential effects of the factors can be ascertained from the present experiment.

Data was placed into vectors of length 440 with a typical case being named WSL0275A. Here the first W denoted a waiting time and occurs in the code for all vectors. (One could also look, say, at delay times. These are the customer's waiting time plus his service time.) SL denotes 'single line queueing discipline' whereas the other discipline, joining the shortest and thus presumably the fastest line, is denoted by FS. The first pair of numbers in the name of the data vector, 02 in this example, represent m, the number of servers, and could be 2,5 or 10. The second set of numbers, here '75', gives the traffic intensity, t, and could be 30, 60 or 75, representing values of t=0.30, t=0.60 and t=0.75. Finally the last letter, here 'A', represents the four cases of distributional assumptions given above with possible values A, B, C or D.

COMPUTING

The data were generated on an IBM 370/3033 computer using a PL/1 program called QSIM. The graphics was done with the experimental APL program GRAFSTAT from IBM Research. The coding of the data vectors explained above made it simple to generate the full screen interfaces for GRAFSTAT, and thus the graphs, under program control. Since many combinations of graphs were tried, this procedure proved very valuable.

The plots were obtained in two ways: one by sending GRAFSTAT output on an IBM 3179 Model G2 terminal to an IBM 7372 plotter, the other by sending GRAFSTAT output on an IBM 3270 AT GX color screen to the same plotter. In both cases the transfer is accomplished by invoking the GDDM4 program on the mainframe.

Both APL and the GRAFSTAT program are ideal for handling a large amount of data like this in a very flexible fashion.

GRAPHICAL OUTPUT AND ANALYSIS

The graphics which were designed to display the output of the simulation experiment can be thought of as a two way layout of two-dimensional graphs, with rows of graphs representing a fixed distributional factor (A, B, C or D) and the columns of graphs representing the 'number of servers' factor. (This was selected as the column factor because it was felt that this factor would have a minor effect on waiting times, although the simulation subsequently showed that this assumption was not correct).

The complete layout of all 12 graphs is not shown; in fact for reasons of space, only the first row of graphs having common distributional factor A will be discussed.

Each graph in the two way layout represents six cases, three traffic intensities and two queueing disciplines (see Figure 1). The two queueing disciplines for each of the three values of traffic intensity are placed close together on the graph since queueing discipline is the primary factor whose effect on the response - waiting time - is to be examined. This 'waiting time' response is represented on the vertical axis by a **quantile plot**.

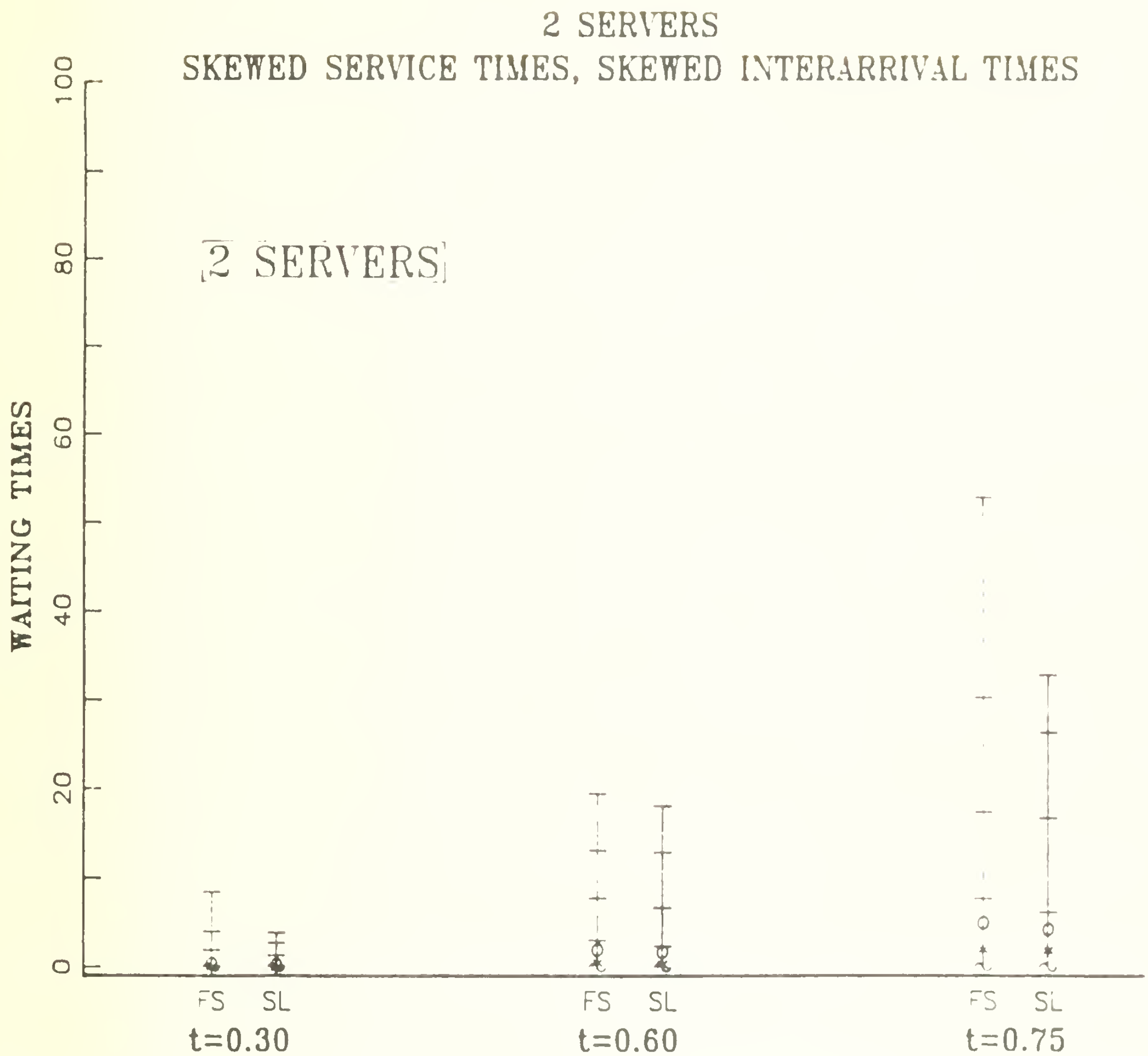

Figure 1. **Two** servers (m=2) and distributional case A - very variable (positively-skewed) service; very variable (positively-skewed) interarrival times. The two quantile plots of the 440 simulated waiting times at each of the three traffic intensities are grouped together and differentiated by color. Blue is used for the 'fastest service', individual choice, case (FS) and red for the single line (SL) case. Note that there are many zero waiting times and that this aspect is not shown in the graph expect in so far as the quantile plots and the means and medians are bunched near zero.

The **quantile plot** is a vertical line with marks representing various estimated quantiles and the estimated mean of the sample of size 440. The clearest case on which to examine these marks is given in Figure 1 for t=0.75 and the blue (FS) line. The tilda at the bottom of the line is the 0.25-quantile or lower quartile of the wating time sample; going upwards, the cross is the 0.50-quantile or median; the circle is the sample mean. - Above this are horizontal bars representing successively the 0.75-quantile or upper-quartile, the 0.95-quantile, the 0.99-quantile, and the maximum value in the sample.

Concentrating on the case illustrated in Figure 1 (m=2; t=0.75; A) it is seen immediately that the single line discipline (SL) in red gives slightly shorter waiting times than the FS case and that the distributions of waiting times are highly, positively skewed. It should be noted that in all cases considered there were many zero waiting times which resulted in a great squashing of the lower quantiles on the graph. In fact a common alternative to the quantile plot is a boxplot, which many people prefer, but in this case it is difficult to plot since many outliers occur at the low end and have the same value. A boxplot results in the pen plotter just cutting a hole in the paper.

Note also that on this graph one cannot 'see' the extent of the probability of zero waiting times, which is an important qualification of the queueing system. Thus, either a table or a graph of this probability should be given separately. The zero-valued waiting times also make it difficult to employ, for example, logarithmic transformations of the data or the scales. This can be obviated by plotting the non-zero waiting times or by looking at delay times (the waiting time plus the service time). However, this latter quantity is not always meaningful in such a simple queue.

Returning to Figure 1 and contrasting the effect of traffic intensity for the two service disciplines with m=2 and distributional case A, we see first that the effect of traffic intensity is marked. This of course is fairly well known. In particular for the Poisson arrival, Exponential service case, the expected waiting time is proportional to $t/(1-t)^2$ (Gaver and Thompson, 1973). However, the waiting times are not affected markedly by the service discipline at any of the traffic intensities considered. In fact, the apparent differences between the FS and the SL cases may not be statistically significant.

Figures 2 and 3 complete the top row of the two way layout of graphs, showing the effect of number of servers, m, on the waiting times for different traffic intensities and different queueing disciplines. A comparison of Figures 1 and 3 reveals very quickly that if the traffic intensity is held constant, then waiting times tend to decrease as the number of servers increase. This is especially true at low traffic intensties. Moreover going from Figure 1 to Figure 3 it becomes clear very rapidly that the effect of the two different queueing disciplines on waiting times becomes marked as the number of servers increases! Thus, using a single line with very variable service and interarrival times is really effective in lowering the waiting times only when the queue has many servers.

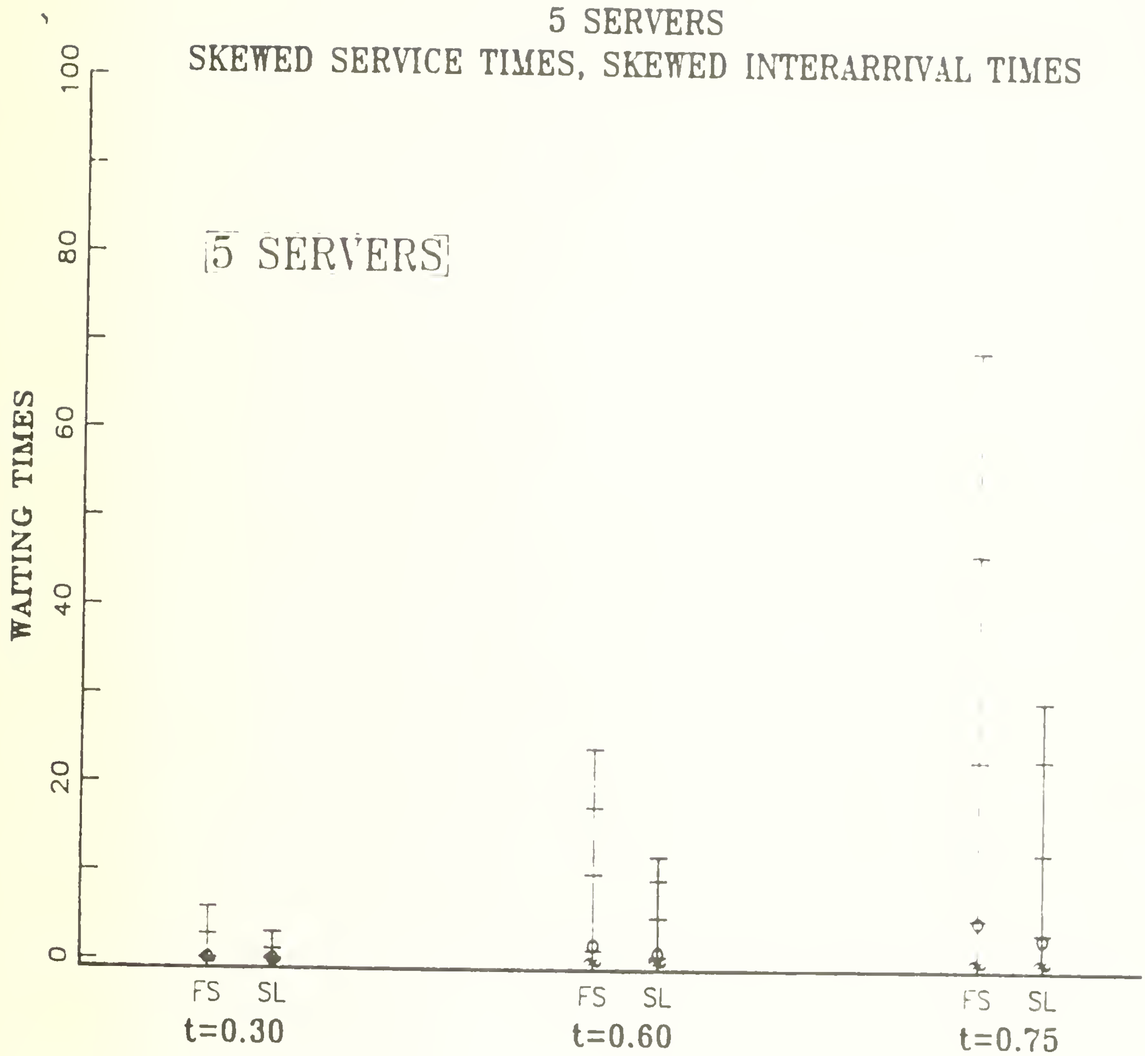

Figure 2. **Five** servers (m=5) and distributional case A - very variable (postively-skewed) service; very variable (positively-skewed) interarrival times. The two quantile plots of the 440 simulated waiting times at each of the three traffic intensities are grouped together and differentiated by color. Blue is used for the 'fastest service', individual choice, case (FS) and red for the single line (SL) case. Note that there are many zero waiting times and that this aspect is not shown in the graph expect in so far as the quantile plots and the means and medians are bunched near zero.

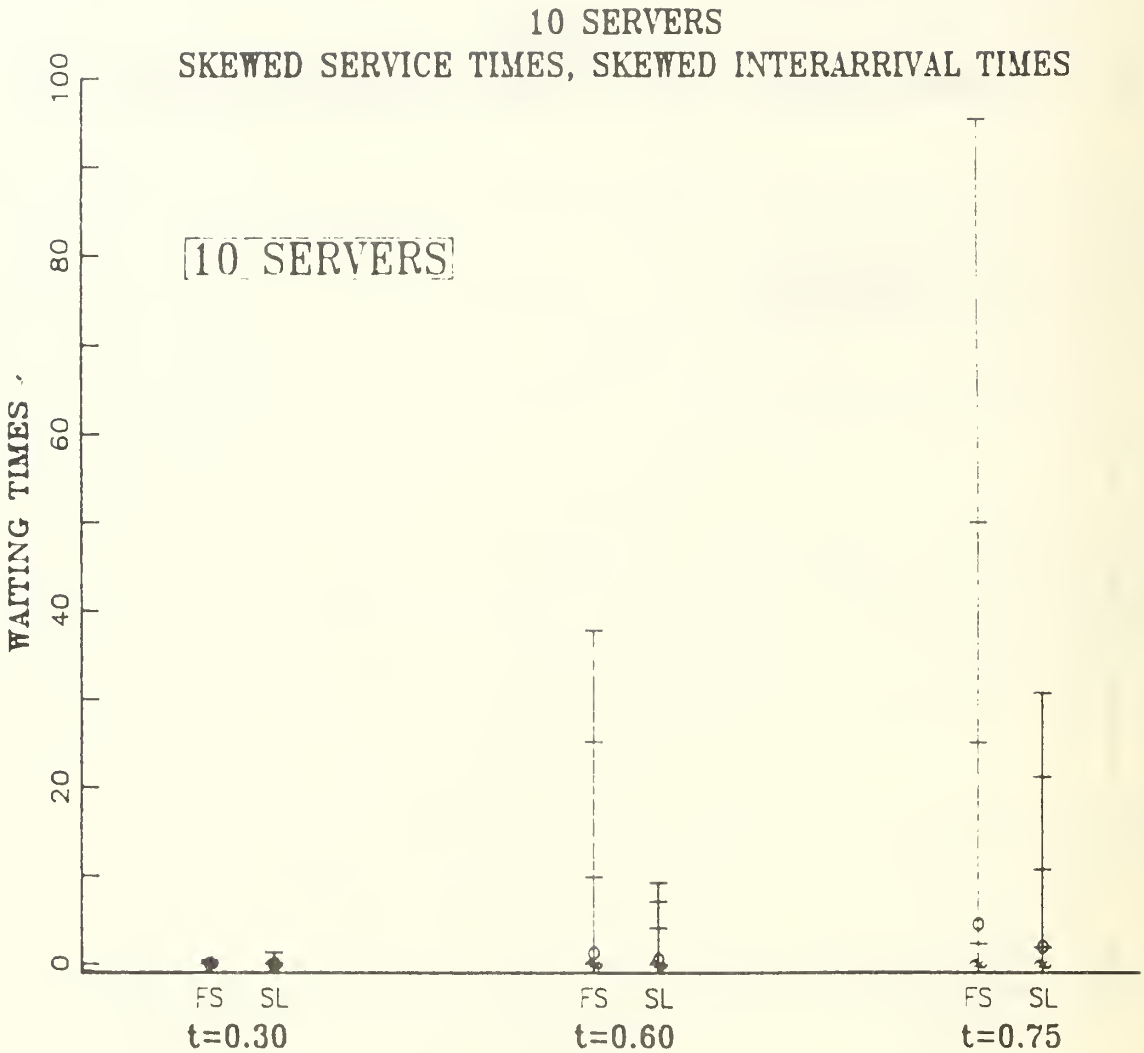

Figure 3. **Ten** servers (m=10) and distributional case A - very variable (positively-skewed) service; very variable (positively-skewed) interarrival times. The two quantile plots of the 440 simulated waiting times at each of the three traffic intensities are grouped together and differentiated by color. Blue is used for the 'fastest service', individual choice, case (FS) and red for the single line (SL) case. Note that there are many zero waiting times and this aspect is not shown in the graph expect in so far as the quantile plots and the means and medians are bunched near zero.

USE OF COLOR AND ENHANCEMENTS

Color has been used to differentiate the closely grouped quantile plots for the two queueing disciplines on the graphs. It is much more effective than, say, line-type in differentiating the two cases to the eye. Actually, a combination of line-type and color seems to work even better but it seems to be a waste of an additional coding device which can be used to further compactify the graphical output. This can be done by superimposing Figure 1 on Figure 2 with a slight left shift and using dashes for the vertical lines in the quantile plots. Similarly Figure 3 is superposed on Figure 2 with the same shift, this time to the right and using dotted lines for the vertical lines in the quantile plots.

If additional colors are available, the coding by line-type can be replaced by coding by color, while the two queueing disciplines are differentiated by line-type. This seems to be the 'best' combination available. However, that many colors are difficult to plot on the available equipment.

CONCLUSIONS

The need to compactify as much as possible the graphical analysis of simulation output so that the results can be encompassed by an analyst is greatly enhanced by the availability of color as a coding device. The types of graphs presented here should be standard tools for simulators and it is hoped that this will be the case when color graphics, and especially the printed version, becomes easier and cheaper to obtain.

ACKNOWLEDGEMENTS

The author is indebted to Dr. P. D. Welch for help with the graphics and for the use of GRAFSTAT. This experimental APL graphics program is available at the Naval Postgraduate School under a test bed agreement with IBM Research.

The author was supported by the Office of Naval Research under Task No. NR 042-469 during the writing of this paper.

REFERENCES

Gaver, D.P. and Thompson, G.L. (1973). Programming and Probability Models in Operations Research. Brooks/Cole: Monterey, California.

Kleinrock, L. (1975). Queueing Systems, Volume 1: Theory. Wiley-Interscience, New York, New York.

DISTRIBUTION LIST

	NO. OF COPIES
Office of Naval Research Arlington, VA 22217	1
Library (Code 0142) Naval Postgraduate School Monterey, CA 93943-5100	2
Office of Research Administration Code 012 Naval Postgraduate School Monterey, CA 93943-5100	1
Professor P. A. W. Lewis Code 55Lw Naval Postgraduate School Monterey, CA 93943-5100	10
Center for Naval Analyses 200 Beauregard Street Alexandria, VA 22311	1
Operations Research Center, Room E40-164 Massachusetts Institute of Technology Attn: R. C. Larson and J. F. Shapiro Cambridge, MA 02139	1
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Library (Code 55) Naval Postgraduate School Monterey, CA 93943-5100	1